DIGITAL INTEGRATION OF SERVICE MODES OF LIBRARIES BASED ON HYBRID METADATA

Fujun Zhang
Library of Shandong University of Science and Technology
Qingdao, Shandong
266590, China

Quanhui Ye
College of Computer Science and Engineering
Shandong University of Science and Technology
Qingdao, Shandong
266590, China

Shuwei Zhang*
College of Electrical and Automation Engineering
Shandong University of Science and Technology
Qingdao, Shandong
266590, China
zhangshuwei1234@yahoo.com

Abstract. Public library is the largest repository of the information society. With the development of the network and the information technology, the demand for digital information is increasing, and library structure and service are beginning to develop in the direction of digital resources. With culture and civilization rooted in various kinds of books and resources, it is of great significance to strengthen the development and construction of cultural resources within libraries, so as to effectively strengthen cultural transmission and promote the integration of cultural resources as well as cultural construction. To explore the service mode of library resources in a digital environment from an integration perspective can effectively promote the construction of national culture. Based on a centralized, distributed and shared metadata management model, this paper designs a hybrid metadata management model suitable for the digital resource management of libraries, and designs a new library digital fusion service platform based on this. At the same time, this paper uses the fuzzy mathematics theory to evaluate the digital fusion service quality of libraries, which provides a new idea for activating the national cultural resources and inheriting civilization, which is of great significance for the promotion of the digital fusion service of libraries.
Keywords: Metadata, library, digital fusion service, resource sharing

Introduction

Libraries being the bearing ground of national civilized memories and playing an irreplaceable role in the real life of mankind, it is extremely important to

∗. Corresponding author

optimize the library management and service model [1-3]. Due to the rapid development of computer network communication technology and other related high and new technologies, information storage, transmission and utilization has undergone tremendous changes. Affected by this, the library collection resources have also begun to change towards digitalization [4]. And the continuous growth of network information resources [5-6] makes it increasingly important to organize and manage these large amounts of information resources effectively. As the data information of data set [7], sound, image, text and other digital information, metadata has become an important tool for the organization and acquisition of digital resources in the modern Internet environment, which plays an important role in the optimization of library service model. For metadata research, the Metadata Consortium has proposed the OIM (Open Information Model) model [8], and the object management organization has proposed the CWM (Public Warehouse Model) model [9].

Zhengwei Sui, Yuan Tian et al. [10] proposed a distributed metadata management framework for mineral information resources with access control through summarizing and evaluating the three typical metadata management modes of centralized, distributed and shared models, taking into account the advantages and disadvantages and the management characteristics of the three models. Based on the advantages of three kinds of metadata, this paper constructs a hybrid metadata management method for the collection, management and dissemination of digital resources in libraries. Chinese scholar Wang Danyang concluded in her work [11] that the innovation of the library digital fusion service could better promote the dissemination and sharing of cultural resources. In this paper, a hybrid metadata management architecture was proposed, a library digital fusion service platform was constructed and the fuzzy mathematics theory was used to evaluate the service quality to make the management of libraries more orderly.

1. The construction of a hybrid metadata management system

Usually, the library service system has a meta-database to manage all the metadata in the library system. There are three kinds of metadata management models, namely, centralized, distributed, shared, which have the same basic features and functions, but also have their own advantages and disadvantages.

1.1 Overview of the three metadata management models

1.1.1 Centralized metadata management

Structure: All tools and database management systems do not require local storage or maintenance of metadata, but have direct access to a unified central knowledge base [12].

Advantages: little resource occupation, low cost, high system resource utilization rate, with only one object storage server in activities. Disadvantages:

When the cluster size is large, all the metadata in the system needs to go through the metadata server which is a centralized access path, and thus the impact on the system performance is large.

1.1.2 Distributed metadata management

Structure: All tools and database management systems are required to store and maintain metadata in their own meta-databases. Advantages: Both remote and local file access can be done through the same system call; the name or address of the server is not part of the file path, and changes to the file storage location do not cause the file path to change; distributed management allows parallel access to metadata to ensure access to metadata. Disadvantages: Due to the complexity of control, the design and implementation of distributed management system is difficult. To solve the problem of consistency maintenance of distributed management, complex protocols such as distributed lock technology are needed, which will cost a lot.

1.1.3 Shared metadata management

Structure: Each tool stores the local metadata in its own knowledge base, then defines the output metadata by defining the corresponding output scheme in the knowledge base, and finally combines the output schemes into a common metadata model. Advantage: Good autonomy, high efficiency of access, good synchronization and consistency, high utilization rate.

Disadvantage: it is difficult to realize in an environment where multiple libraries work together.

1.2 Mixed metadata management methods

Through the analysis, this paper takes the advantages of the above three methods and combines them into a mixed metadata structure to build a hybrid metadata management method, as shown in figure 1.

In the process of building, due to the poor practicality of the centralized model, it is added to the data access layer. In addition, a shared knowledge base is introduced to manage metadata associated with the whole situation and connected to specific components. In order to better control the flow of metadata, the building uses models, ETL and data access tools to better capture and record objects created during data analysis.

2. Construction and evaluation of digital fusion service platform for library resources The construction of the hybrid metadata management system and the aggregation of the library resources [12] laid the foundation for the construction of the library digital fusion service platform. The integration of a library's digital resource service models does not simply refers to the combination of the digital resources and service forms of the library but also includes the innovation of the service forms. Through the research of library digital fusion service model level,

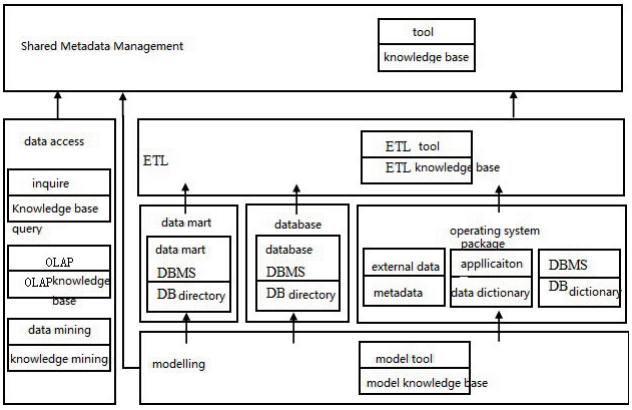

Figure 1: Mixed metadata management building results

the design of platform and the comprehensive evaluation of influence factors of fusion service based on user demand, this paper realizes the construction of the digital fusion service model of the whole library resources.

1.3 Digital resource fusion service platform construction

Through the construction of a unified portal platform, this study provides users with one-stop retrieval and evaluation services [13-1], with its basic framework divided into four levels, as shown in figure 2:

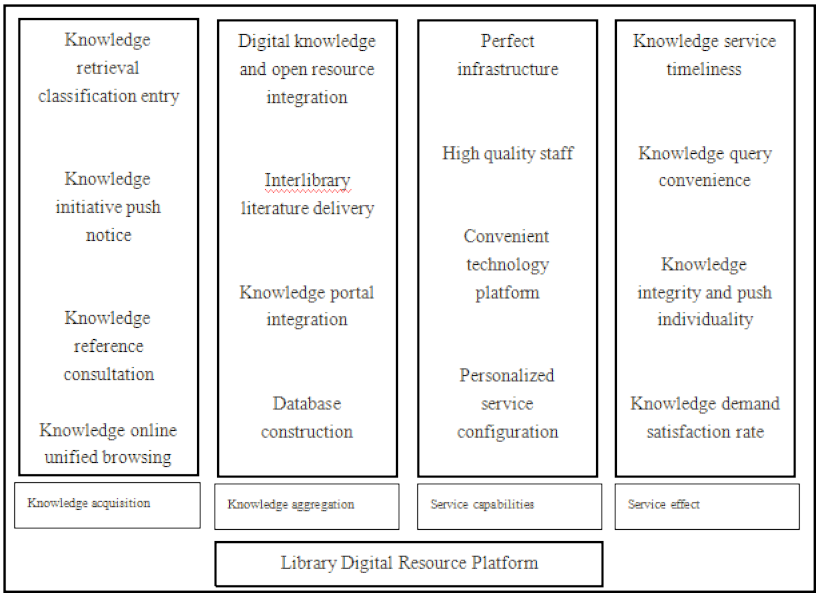

Figure 2: Library resources digital convergence service platform framework

The first layer is the knowledge acquisition layer. Metadata mapping, RDF or metadata program are used to achieve resource integration and build the linkage between the information so as to provide users with the most basic fusion search, online browsing and push services; the second layer is the knowledge aggregation layer. Through the associated data and other technologies, the library digital resources and portal resources are integrated to constitute a resource knowledge network; the third layer is the innovative service capability layer. With perfect infrastructure, high-quality staff, convenient technology platform and personalized service configuration, users are provided with the best services; the fourth layer is the service effect layer. The effectiveness of digital resources service effects are judged by knowledge service timeliness, knowledge query convenience, knowledge provision integrity and push individualization and knowledge requirements satisfaction rate.

1.4 Evaluation of library digital fusion service quality with fuzzy mathematical model

1.4.1 Establishment of evaluation model

In this study, V is set as the evaluation set, U as the factor set: $V=\{v_1, v_2, \ldots, v_n\}$, $U = \{u_1, u_2, \ldots, u_n\}$; u_i is the main factor affecting the integration services of the assessed digital library; v_j refers to rating levels. Since the membership degree of u_i for the evaluation set is a_{ij}, the result of the evaluation of the ith factor u_i is obtained and a single factor evaluation matrix A is constructed.

$$A = \begin{bmatrix} 1 & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & 1 \end{bmatrix}.$$

In essence, through the number in the $[0, 1]$ interval, A expresses the fuzzy relation of the factor set U to the evaluation set V, reflecting the degree of fuzzy relationship between the two. U, V, A forms the fuzzy comprehensive evaluation model for evaluating the quality of digital library integration services. We divided the reviews into five grades, namely:

$$V = \{\text{excellent}\,(v_1), \text{good}\,(v_2), \text{average}\,(v_3), \text{qualified}(v_4), \text{unquallified}(v_5)\}.$$

The digital library integration service quality evaluation (U) can be divided into four first-level indicators, which are knowledge acquisition ($S = u_1$), knowledge aggregation ($X = u_2$), service capabilities ($Y = u_3$), service effectiveness ($Z = u_4$).

Each first level indicator includes four secondary indicators. Knowledge acquisition (S) includes: knowledge retrieval and classification entries (S1), knowledge active push (S2), knowledge reference (S3), knowledge online browsing (S4). Knowledge integration (X) includes: digital knowledge and open resource

integration (X1), interlibrary literature transmission (X2), knowledge portal integration (X3), database construction (X4). Service capability (Y) includes: perfect infrastructure (Y1), high-quality staff (Y2), convenient technology platform (Y3), personalized service configuration (Y4). The service effect (Z) includes: knowledge service timeliness (Z1), knowledge query convenience (Z2), knowledge provision integrity and personalized push (Z3), knowledge demand satisfaction rate (Z4).

Since the focus of evaluation of each factor in U, i.e., u_i, is different, the focus of its weight is different. By determining the relative importance and weight of each factor, the weight vector $\omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ is determined, where $0 \leq \omega_1 \leq 1$ and $\sum_{i=1}^{n} \omega_i = 1$. Then, the fuzzy comprehensive evaluation value B is determined, i.e., $B = \omega OA,$, where O refers to a generalized synthetic operation.

1.4.2 Consistency test

The judgment matrix has the following properties:

$a_{ii} = 1$;

$a_{ij} = 1/a_{ji}$;

$a_{ij} = a_{ik}/a_{jk}$; $(i, j, k = 1, 2, \ldots, n)$.

As long as the a_{ij} in the judgment matrix satisfies the above three relations, the judgment matrix has complete consistency. When $a_{ij} > 0$, A is a positive reciprocal matrix.

The largest eigenvalue of A is a positive real number and the component of the corresponding eigenvector is the same number. If the unit eigenvector corresponding to the largest eigenvalue $\lambda_{\max}$ is $W = (w_1, \ldots, w_n)^T$, then $a_{ij} = w_i/w_j (i, j = 1m2m \ldots, n)$, $\forall i, j = 1, 2 \ldots, n$. $W = (w_1, \ldots, w_n)^T$ is the weight vector of each indicator line for the superior indexes in the same subset.

However, since each factor in A is obtained through the pairwise comparison of indicators in the same indicator set, it is not necessarily credible to take the positive unit eigenvector corresponding to the maximum eigenvalue $\lambda_{\max}$ of matrix A_{as} the weight vector. Therefore, a consistency test is required.

The consistency test indicator used in this paper is $C.R. = \frac{C.I.}{R.I.}$, where $C.I. = \frac{\lambda_{\max} - n}{n-1}$, $\lambda_{\max}$ refers to the maximum eigenvalue of the judgment matrix and n is the order of the judgment matrix.

The consistency of the matrix is determined by the size of the value of C.I. . The greater the value of C.I., the greater the degree of deviation of the judgment matrix from complete consistency; the smaller the value of C.I., the closer the judgment matrix is to complete consistency. At the same time, the larger the order number n of the judgment matrix, the greater the value of the deviation from the complete consistency index C.I.; the smaller the order number n, the smaller the value of the deviation. When $n < 3$, the judgment matrix is always completely consistent. The ratio of the consistency index C.I. of the judgment

matrix to the average random consistency index R.I. of the same order is the random consistency ratioC.R..

1.4.3 Examples

$U = \{$knowledge acquisition $(u1)$, knowledge aggregation $(u2)$, service capability $(u3)$, service effect $(u_4)\}$, i.e., $U = \{u_1, u_2, u_3, u_4\}$. The result comparison matrix for each factor is as follows:

$$Q = \begin{bmatrix} 1 & 1/3 & 1/2 & 1/4 \\ 2 & 1 & 3 & 1 \\ 2 & 1/2 & 1 & 1/2 \\ 4 & 1 & 2 & 1 \end{bmatrix}$$

$$\begin{cases} Q\omega == \lambda_{\max}\omega \\ \sum_{i=1}^{n} \omega_i = 1. \end{cases}$$

Then, the weight vector $\omega = \{0.15, 0.29, 0.21, 0.34\}$ is obtained.

Through investigation, it is found that the proportion of excellent, good, average, qualified and unqualified is $30\%, 20\%, 30\%, 10\%$ and 10%. Hence, the single factor evaluation A_1 on u_1 level is obtained, as does the other factor evaluation. Afterwards, a fuzzy evaluation matrix A is obtained, as follows:

$$A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix}, \text{ assume} A = \begin{bmatrix} 0.3 & 0.2 & 0.3 & 0.1 & 0.1 \\ 0.3 & 0.3 & 0.2 & 0.2 & 0 \\ 0.3 & 0.4 & 0.2 & 0.1 & 0 \end{bmatrix} \begin{matrix} \\ 0.2 & 0.2 & 0.4 & 0.2 & 0.1 \\ \\ \end{matrix}.$$

Then, the following can be obtained:

$$B = \omega oA = (0.15, 0.29, 0.21, 0.35) \circ \begin{bmatrix} 0.3 & 0.2 & 0.3 & 0.1 & 0.1 \\ 0. & 0.3 & 0.2 & 0.2 & 0 \\ 0.2 & 0.2 & 0.4 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.2 & 0.1 & 0 \end{bmatrix}$$
$$= (0.2790, 0.2990, 0.2570, 0.1500, 0.0360).$$

After normalization, $B' = 0.273, 0.293, 0.252, 0.147, 0.035$ can be obtained.

It can be concluded that 27.3% of users rated as excellent, 29.3% of users rated as good, 25.2% of user rated as average, 14.7% of users rated as qualified while 3.5% of users considered the quality poor, i.e., unqualified. The results of the evaluation applying fuzzy comprehensive evaluation method have strong credibility and comparability. The overall service quality of the digital library integration service platform is mostly excellent, good and average, and we can judge that the platform is satisfactory to customers, though, improvement exist.

2. Conclusion

With the advent of the large data era and the rise of Internet of things, cloud computing and mobile Internet, the development of digital library integration services is becoming an inevitable trend, with its construction becoming more and more challenging. The importance of book resources has led the country to gradually attach importance to and develop the library service management. This paper studied the digital fusion service model of libraries based on hybrid metadata in an environment with rapid information technological development and state's support to studies of library service models. Libraries should correctly establish their own service goals, and their digital integration services not only need to provide information and knowledge resources, but also need to explore the deep value of digital cultural resources. In this paper, a library digital fusion service platform was built and a comprehensive evaluation of library fusion service quality applying the fuzzy mathematical model was carried out, though, there are still shortcomings. Particularly, the validity detection of mixed metadata and the practicality detection of the metadata management scheme need to be further studied.

References

[1] D.B. Lindenmayer, M.P. Piggott, B.A. Wintle, *Counting the books while the library burns: why conservation monitoring programs need a plan for action* [J]. Frontiers in Ecology & the Environment, 11 (2013), 549-555.

[2] C.J. Chou, C.W. Chen, C. Conley, *A systematic approach to generate service model for sustainability [J]*, Journal of Cleaner Production, 2930 (2012), 173-187.

[3] S.S.C. Shang, E.Y. Li, Y.L. Wu, et al., *Understanding Web 2.0 service models: A knowledge-creating perspective [J]*, Information & Management, 48 (2011), 178-184.

[4] C.L. Anderson, R. Agarwal, *The Digitization of Healthcare: Boundary Risks, Emotion, and Consumer Willingness to Disclose Personal Health Information* [J]. Information Systems Research, 22 (2011), 469-490.

[5] J.Y. Wang, Z. Zhu, *Integration system of network information resources based on multi-agent collaboration*,[C] Fuzzy Systems and Knowledge Discovery (FSKD), 2011 Eighth International Conference on IEEE, 2011, 2044-2049.

[6] X. Zhu, *The development and utilization research of network information resources for English learning* [C] Advanced Research and Technology in Industry Applications, IEEE, 2014, 138-140.

[7] R.J. Barro, J.W. Lee, *A New Data Set of Educational Attainment in the World*, 1950-2010. NBER Working Paper No. 15902 [J]. National Bureau of Economic Research, 104 (2010), 184198.

[8] M.S. Song, M.S. Kim, S.H. Lee, *Application of Open Information Model for the Information Management on Building Flood Damage* [J] 27 (2014), 565-572.

[9] G. Miao, J. Gou, *Yearbook Data Integration Based on Common Warehouse Model* [C] Special Session on Project Management and Service, Science, 2011, 569-573.

[10] Z. Sui, Y. Tian, X. Wang, et al., *A framework for distributed metadata management of mineral information resources with access control* [C]// International Conference on Geoinformatics, 2013, 1-4.

[11] Dan-Yang Wang, *The Dimension Design and Cooperation Mode on the Realization of the Integration of the Digitizing Library Archives and Services* [J] Journal of Tangshan Teachers College, 38 (2016), 140-142.

[12] E. Selberg, O. Etzioni, *The MetaCrawler architecture for resource aggregation on the Web* [J] IEEE Expert, 2010, 12 (2010), 11-14.

[13] N. Silvester, B. Alako, C. Amid, et al., *Content discovery and retrieval services at the European Nucleotide Archive*, [J] Nucleic Acids Research, 2015, 43 (Database issue), 23-9.

[14] M. Goto, J. Ogata, *PodCastle: Recent Advances of a Spoken Document Retrieval Service Improved by Anonymous User Contributions*, [C] INTERSPEECH 2011, Conference of the International Speech Communication Association, Florence, Italy, August, 2011, 104-113.

Accepted: 25.03.2017